Data S1 - Online supporting information for *Occupancy Modeling Species-Environment Relationships with Non-ignorable Sampling Designs*

Kathryn M. Irvine, Thomas J. Rodhouse, Wilson J. Wright, Anthony R. Olsen

Data S1: R-code to recreate empirical example results using pseudo-likelihood and likelihood estimation for single- season occupancy models.

This Appendix includes the R code used for recreating the results in Figure 2. These files require the dplyr and ggplot2 packages.

1. OR_model_fits.R:

   This file fits the single-season occupancy model using P-ML and ML for all three species and returns Figure 2 in the main text.

2. or_covs.csv:

   contains the empirical example Oregon dataset with columns for CONUS_10KM (= sample unit ID), Elevation_mean, and Forest_Percent_Gap for all sample units in Oregon ($N = 2660$). Elevation_mean was based on US Geological Survey 10-m digital elevation model and summarized to the 10-km x 10-km sample unit (https://lta.cr.usgs.gov/NED). Forest_Percent_Gap was created from GAP land cover 30-m resolution land cover map

(https://gapanalysis.usgs.gov/gaplandcover/). Forest cover was based on aggregated 'forest and woodland systems' and summarized to each 10-km x 10-km sample unit.

3. lano_dets.csv:

   Sample Unit ID is the same as CONUS_10KM in or_covs.csv. Each row is the observed detection/non-detection history for silver-haired bat (LANO; *Lasionycteris noctivagans*) at each of 91 sample units in 2016.

4. mylu_dets.csv:

   Sample Unit ID is the same as CONUS_10KM in or_covs.csv. Each row is the observed detection/non-detection history for little brown myotis (MYLU; *Myotis lucifugus*) at each of 91 sample units in 2016.

5. myvo_dets.csv:

   Sample Unit ID is the same as CONUS_10KM in or_covs.csv. Each row is the observed detection/non-detection history for long-legged myotis (MYVO; *Myotis volans*) at each of 91 sample units in 2016. The column SurveyType is the strata membership for the 91 surveyed sample units in 2016: "Prob" for the NABat Oregon sample units, "nonprob" for the legacy sites, or "FWS" for the NABat FWS R1 sample units.

6. pmle_functions.R:

   This file is the required code for implementing the models called in the file "OR_model_fits.R." It contains two functions:

   - *logL.fun*: log-likelihood of a proposed occupancy model with sample weights
   - *occ_pmle*: fits the single-season occupancy model with required inputs for occupancy model; detection model; detection history matrix, NAs okay; dataframe of site-level covariates; named list of visit-level covariates; vector of weights for PMLE fits.

     Example call to function to estimate occupancy model using P-MLE with forest and elevation occupancy covariates and constant detection for MYVO.

```
occ_pmle(~forest+elev.m,
                        ~1,
                        myvo_dh,
                        myvo_cov2,
                        list(blank=myvo_dh),
                        weights = myvo_wts)
```

Returns the following output (example for MYVO P-MLE results):

```
$params
[1] "alpha[0]" "beta[0]"  "beta[1]"  "beta[2]"


$estimates
[1] -0.72918766  2.73029511  2.24911931  0.01662806


$std.errors
[1] 0.1527564 1.3974652 1.1977440 0.4570784


$hessian
           [,1]      [,2]      [,3]       [,4]
[1,] 50.5219064  5.790555 -5.092232  0.1891535
[2,]  5.7905550  5.897079 -6.411697  0.2035330
[3,] -5.0922315 -6.411697  7.845033 -1.1351340
[4,]  0.1891535  0.203533 -1.135134  5.7853115


$converge
[1] 0


$AIC
```

```
[1] 372.8027
```